

Detecting atomic clock frequency trends using an optimal stopping method

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2016 Metrologia 53 S89

(<http://iopscience.iop.org/0026-1394/53/3/S89>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 95.236.229.60

This content was downloaded on 23/05/2016 at 10:25

Please note that [terms and conditions apply](#).

Detecting atomic clock frequency trends using an optimal stopping method

C Zucca¹, P Tavella² and G Peskir³

¹ Department of Mathematics ‘G. Peano’, University of Torino, Via Carlo Alberto 10, 10123 Turin, Italy

² Istituto Nazionale di Ricerca Metrologica (INRiM), Strada delle Cacce 91, 10135 Turin, Italy

³ School of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

E-mail: cristina.zucca@unito.it, tavella@inrim.it and goran@maths.manchester.ac.uk

Received 18 February 2016, revised 28 March 2016

Accepted for publication 29 March 2016

Published 20 May 2016



Abstract

We consider the problem of rapid detection of anomalies in atomic clocks. We assume that the clock error evolution can be modeled by a Wiener process with a drift changing from 0 to $\mu \neq 0$ at some random unobservable time θ and we approach the problem by applying the quickest detection method. We describe the scope and the limits of the method and we apply it to clock experimental data.

Keywords: atomic clock, frequency jumps, detection, optimal stopping

(Some figures may appear in colour only in the online journal)

1. Introduction

Recently, particularly in space applications such as the global satellite navigation systems (GNSS), anomalous clock behavior has been experimentally observed: phase or frequency jumps, increase of instability, general non-stationary trends. Such anomalous behavior has an important impact on the global system and may affect its overall performance very seriously, as it compromises the reliability of the atomic clocks which is a crucial point in navigation systems.

When a clock has an anomaly, this has to be detected very quickly and the users have to be informed promptly to avoid the use of a failing clock and hence a wrong position estimate. Rapid detection of anomalies is therefore of the utmost importance together with the need for very low false alarm probability.

Here we approach the problem by applying the quickest detection method (see [1, 2]). We assume that the clock deviation evolution can be modeled by a Wiener process $X = (X_t)_{t \geq 0}$ and that we observe a trajectory with a drift changing from 0 to $\mu \neq 0$ at some random time θ (see [3, 4]). Our task is to find a stopping time τ of the process X that is as close as possible to the unknown time θ .

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

We describe the scope and the limits of the method applied to the detection of the atomic clock anomalies and we present results on the application of this method to clock experimental data.

2. Detection method

Suppose that we observe a trajectory of a Wiener process $X = (X_t)_{t \geq 0}$ with a drift changing from 0 to $\mu \neq 0$ at some random time θ . We assume that θ is a random variable that takes the value 0 with probability π and it is exponentially distributed with parameter $\lambda > 0$ given that $\theta > 0$. We are looking for a stopping time τ of the observed process X that is as close as possible to the unknown time θ .

The observed process X can be described as

$$\begin{aligned} dX_t &= \mu I(t \geq \theta) dt + \sigma dW_t \\ X_0 &= 0 \end{aligned} \quad (1)$$

which is equivalent to

$$X_t = \begin{cases} \sigma W_t & t < \theta \\ \mu(t - \theta) + \sigma W_t & t \geq \theta \end{cases} \quad (2)$$

where $W = (W_t)_{t \geq 0}$ is a standard Wiener process, $\mu \neq 0$ and $\sigma^2 > 0$ are given and fixed constants, and $I(\cdot)$ is the indicator function.

The aim is to find a stopping time of the process X that is as close as possible to the unknown time θ by avoiding false alarms and minimizing the detection delay. This means that we are looking for a stopping time τ that minimizes the risk function

$$V(\pi) = \inf_{\tau} [\mathbb{P}_{\pi}(\tau < \theta) + c\mathbb{E}_{\pi}(\tau - \theta)^+] \quad (3)$$

where $\mathbb{P}_{\pi}(\tau < \theta) =: PFA$ is the probability of false alarm, i.e. the probability of declaring the alarm when it is not needed; $\mathbb{E}_{\pi}(\tau - \theta)^+ =: D$ is the expected detection delay, i.e. the mean delay to detect the disorder, and c is a positive constant that balances the two quantities. By $(z)^+$ we denote the positive part of z (which equals z when positive and 0 otherwise).

It is known (see [1, 2]) that the problem (3) can be rewritten as an optimal stopping problem with the risk function

$$V(\pi) = \inf_{\tau} \mathbb{E}_{\pi} \left[1 - \Pi_{\tau} + c \int_0^{\tau} \Pi_t dt \right] \quad (4)$$

expressed in terms of the *a posteriori* probability process $\Pi_t = \mathbb{P}_{\pi}(\theta \leq t | \mathcal{F}_t^X)$ for $t \geq 0$, where $\mathbb{P}_{\pi}(\Pi_0 = \pi) = 1$ and $(\mathcal{F}_t^X)_{t \geq 0}$ is the natural filtration of the process X , which means the history of X up to time t for $t \geq 0$. Observe that Π_t takes values in $[0, 1]$ for all $t \geq 0$.

The following theorem (see [1, 2]) presents a solution to the optimal stopping problem.

Theorem 2.1. *The value function $V(\pi)$ is given by*

$$V(\pi) = \begin{cases} (1 - A) + \int_A^{\pi} \psi(\rho) d\rho & \text{if } \pi \in [0, A) \\ 1 - \pi & \text{if } \pi \in [A, 1] \end{cases} \quad (5)$$

for $\pi \in [0, 1]$, where A is the unique root of the equation

$$\psi(A) = -1 \quad (6)$$

and

$$\psi(\pi) = -\frac{c}{\gamma} e^{-\frac{\lambda}{\gamma} \alpha(\pi)} \int_0^{\pi} \frac{e^{\frac{\lambda}{\gamma} \alpha(\rho)}}{\rho(1 - \rho)^2} d\rho \quad (7)$$

with $\alpha(\pi) = \log(\frac{\pi}{1 - \pi}) - \frac{1}{\pi}$ and $\gamma = \frac{\mu^2}{2\sigma^2}$.

The stopping time

$$\tau_A = \inf\{t \geq 0 : \Pi_t \geq A\} \quad (8)$$

is optimal in $V(\pi)$.

It is known (see [1, 2]) that the *a posteriori* probability process $\Pi = (\Pi_t)_{t \geq 0}$ can be computed as

$$\Pi_t = \frac{\Phi_t}{1 + \Phi_t} \quad (9)$$

where

$$\Phi_t = e^{Y_t} \left(\frac{\pi}{1 - \pi} + \lambda \int_0^t e^{-Y_s} ds \right) \quad (10)$$

and

$$Y_t = \lambda t + \frac{\mu}{\sigma^2} \left(X_t - \frac{\mu}{2} t \right). \quad (11)$$

Remark 2.1. From (9)–(11) we note that the random variable Π_t requires knowledge of the process X up to time t only. This means that the corresponding detection algorithm is a real time algorithm.

Remark 2.2. We emphasize that, in order to apply the theorem, the following assumptions need to be satisfied:

- the underlying observed process is a Wiener process which changes its drift at a random time θ ;
- the new drift μ and the diffusion coefficient σ of the Wiener process are known;
- the disorder time θ is exponentially distributed and its parameter λ is known;
- the constant c (or, equivalently, the false alarm probability $1 - A$) is known;
- π , i.e. the probability that the disorder time θ has occurred at time zero is known.

2.1. The algorithm

Here we explain, step by step, the detection algorithm.

- Hypothesize that the process X and its parameters μ , σ , λ , π are given.
- Now there are two possible approaches: we can first fix either the parameter c or the parameter A . In the first case, we fix c and compute A as the unique root of the equation (7). In the second case, we fix A and, recalling that $\psi(A) = cf(A)$ is a linear function of c , we find c using (7).
- We can then apply the detection algorithm, i.e. we are waiting for the first time τ_A such that the *a posteriori* probability process Π defined in (9) hits the boundary A (assuming that $\pi < A$). Note that we can build Π_t in (8) using (9)–(11). Moreover, to compute the integral in (10) it is useful to make use of the numerical rectangle method [5].

Remark 2.3. The second case of step (ii) is more useful because the parameter A has a probabilistic interpretation. Indeed, the quantity $1 - A$ is the probability of false alarm and often, in applications, it is very important that this probability is sufficiently close to zero. For this reason, in what follows, we will assume that the value of A is given and fixed.

3. Expected delay and probability of false alarm

In assessing the performance of the method it is important to study the delay in detection, i.e. $(\tau_A - \theta)^+$. It is a random variable whose characteristics are not described in the literature and we note that it is possible to write its expected value D in closed form.

We first need to express the function ψ in closed form. Computing explicitly the integral in (7), we get

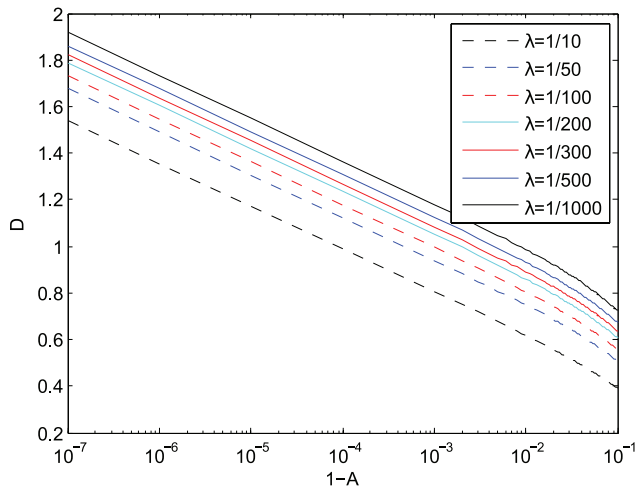


Figure 1. Expected detection delay D as a function of the $PFA = 1 - A$ with $\mu = 5$, $\sigma = 1$ and several values of λ .

$$\psi(x) = -\frac{ca}{\lambda} \left[\frac{a^a \exp\left(\frac{a(1-x)}{x}\right)}{a+1} \left(\frac{x}{1-x}\right)^{-a} \Gamma\left(-a, a\left(\frac{1-x}{x}\right)\right) + \frac{1}{a+1} \left(\frac{x}{1-x}\right) \right] \quad (12)$$

where $a = \frac{\lambda}{\gamma}$ and $\Gamma(\cdot, \cdot)$ is the incomplete gamma function [6]. Using the previous result the expected delay can be written as

$$D = \mathbb{E}(\tau_A - \theta)^+ = -\frac{1}{c} \int_{\pi}^A \psi(x) dx = \frac{a}{\lambda(a+1)} [(\pi + \log(1-\pi)) - (A + \log(1-A))] + \frac{a^{a+1}}{\lambda(a+1)} \int_{\frac{1-A}{A}}^{\frac{1-\pi}{\pi}} \frac{\Gamma(-a, ay) y^a e^{ay}}{(y+1)^2} dy. \quad (13)$$

Note that the expected delay depends on the process parameters μ and σ just through the *signal-to-noise ratio* μ/σ . This quantity is very important for the reliability of the algorithm as shown in the following examples.

Here we analyse the behavior of the expected delay as a function of the probability of false alarm $PFA = \mathbb{P}_{\pi}(\tau_A < \theta) = 1 - A$ as the parameters λ and μ/σ change.

Figures 1 and 2 show the expected delay D as a function of $PFA = 1 - A$ and the parameter λ with $\mu = 5$, $\sigma = 1$. We can see that, fixing PFA , as λ decreases, the expected delay increases. Indeed, a small λ means that the expected waiting time for the change of drift to occur is large and it implies a higher expected delay in the detection, as the method is more cautious in declaring an anomaly which is known to be rare. Nevertheless, even with very small PFA , the expected delay is just one time unit.

Figure 3 shows the expected delay D as a function of the $PFA = 1 - A$ and the signal-to-noise ratio μ/σ with $\lambda = 1/360$. We can see that, fixing PFA , as μ/σ decreases, the expected delay increases. Actually, a small signal-to-noise ratio makes a disorder time more difficult to detect and gives a longer expected delay. If $\mu/\sigma > 1$ the detection algorithm is reliable,

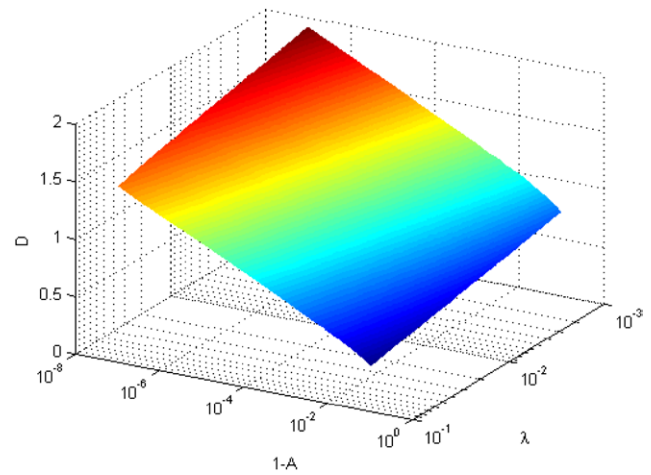


Figure 2. 3D representation of the expected detection delay D as a function of the $PFA = 1 - A$ and λ with $\mu = 5$ and $\sigma = 1$.

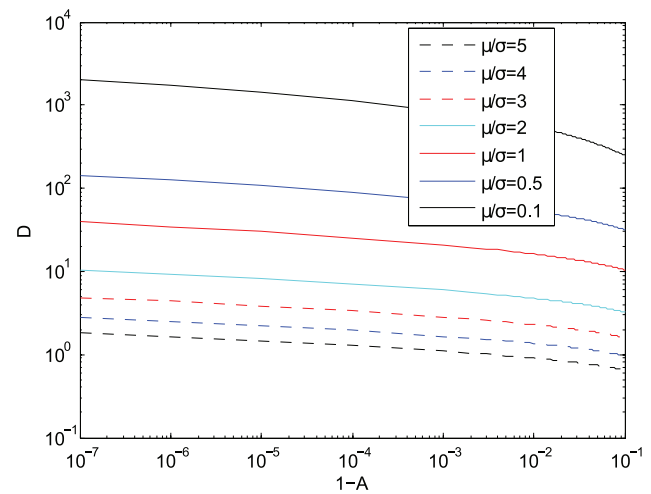


Figure 3. Expected detection delay D as a function of the $PFA = 1 - A$ with $\lambda = 1/360$ and several values of the signal-to-noise ratio μ/σ .

while if $\mu/\sigma \leq 1$ and gets closer to zero, the expected delay grows significantly and the result becomes less applicable unless the PFA is enlarged as shown in the application to data from the New Horizons Project [7].

4. Simulations

In order to illustrate the behavior, the performance and the dependence on parameters of the detection algorithm, in this section we present some simulated examples.

We have already pointed out that the algorithm requires the knowledge of parameters μ , σ , λ , π but, in applications, their values are not known and should be estimated:

- the parameter λ may be estimated from previous observations of the drift change of the process;
- the diffusion coefficient σ can be estimated using the Allan variance [8] or the increments of the process X ;
- $\pi = \mathbb{P}(\theta = 0)$ is fixed in advance by the application problem. It is often set to zero and it represents the probability that the anomaly occurs at time zero.

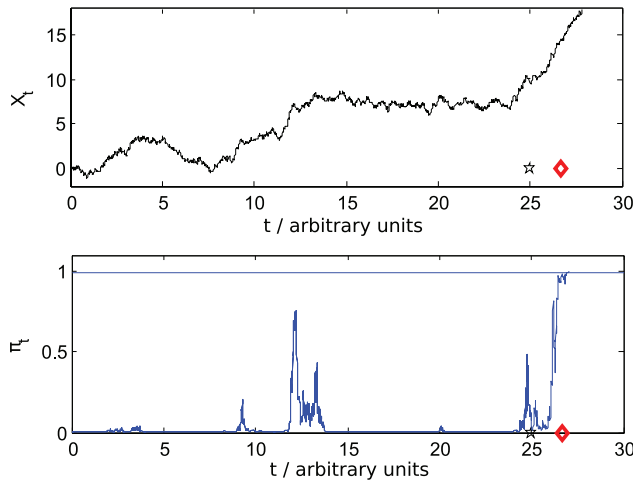


Figure 4. Sample path of the observed process X with parameters $\mu = 3$, $\sigma = 1$ and the corresponding *a posteriori* probability process Π . The disorder time $\theta = 25$ is indicated by a star, while the optimal stopping time $\tau_A = 26.63$ is indicated by a diamond for $1 - A = 0.01$.

- the parameter μ is the critical point in the problem. If the data are analyzed *a posteriori* it is possible to estimate its value using the sample mean of the increments but the reliability of this estimate depends on the signal-to-noise ratio. Moreover, it is well known from the statistics literature (dating back to 1940–50s) that it is impossible to estimate it properly without a huge number of observations. Running the algorithm in real time we may not know the new drift value in the incoming change and we may need to guess this value from previous experience.

Hereafter we study the robustness of the method, varying μ and σ , through simulations.

Figure 4 shows a sample path of the observed process X with parameters $\mu = 3$, $\sigma = 1$, $\pi = 0$ and the corresponding *a posteriori* probability process Π with $PFA = 1 - A = 0.01$. The disorder time is $\theta = 25$ and the algorithm recognizes $\tau_A = 26.63$ as the optimal stopping time. From the figure we can see that, as the trajectory of X increases, the trajectory of Π increases towards the threshold A .

One of the main problems in the application of the detection algorithm is the knowledge of the new drift μ . Indeed, this parameter is often not known.

Figure 5 shows a simulated sample path of the process X with parameters $\mu = 3$, $\sigma = 1$, $\lambda = 1/360$, $\pi = 0$ and the corresponding process Π with $A = 0.97$ and different guessed values of the drift μ . Table 1 lists the values of τ_A and the theoretical expected detection delay for different values of μ .

In this case we are in a favorable situation as the signal-to-noise ratio is greater or equal than 1. Indeed, we can see that the disorder time $\theta = 300$ is recognized by the algorithm even if we set a wrong μ in the application of the method. We have to note that we only apply the method with a ‘slightly’ wrong value of μ , as at least the order of its magnitude should be known.

Next we test the detection algorithm’s response when a wrong choice of λ has been made.

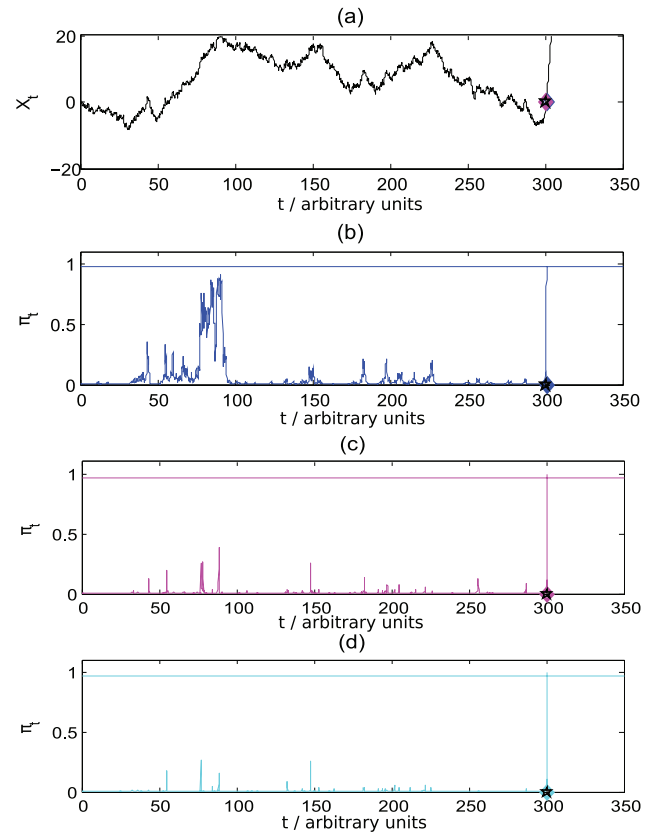


Figure 5. Sample path of the process X with parameters $\mu = 3$, $\sigma = 1$ and the corresponding *a posteriori* probability process Π with $\mu = 1$ (blue), $\mu = 3$ (magenta), $\mu = 5$ (cyan). The disorder time θ is indicated by a star, while the optimal stopping times τ_A are indicated by a diamond for $1 - A = 0.03$.

Table 1. Detection times of the simulated trajectory from figure 5 and the theoretical expected detection delay for different values of μ .

μ	τ_A	D
1	301.25	13.72
3	300.65	2.00
5	300.65	0.80

Note: The disorder time θ equals 300 in this case.

Figure 6 shows a simulated sample path of the process X with parameters $\mu = 3$, $\sigma = 1$, $\lambda = 1/360$, $\pi = 0$ and the corresponding process Π with $A = 0.97$ and different values of λ . Table 2 lists the values of τ_A and the theoretical expected detection delay for different values of λ .

We can see that the disorder time $\theta = 300$ is recognized by the algorithm even if we set a wrong λ .

5. Real data

In this section we test the algorithm on some real clock data [9]. For application reasons in the following we set $PFA = 1 - A = 10^{-7}$.

The first data set was extracted by a GLONASS space caesium clock. The sample size of the data set is 8064 and the sampling time is 300 s. The time deviation data and the increments are represented in figures 7(a) and (b). Analyzing

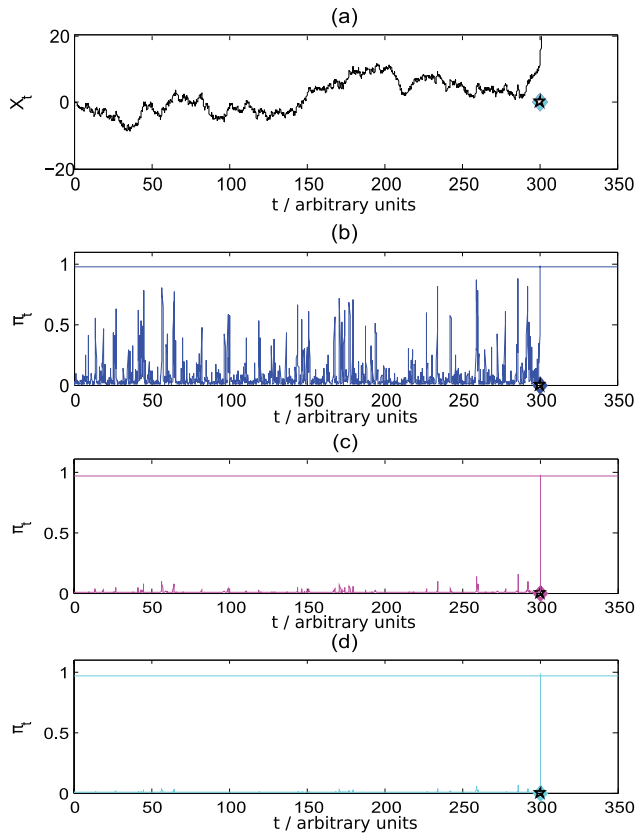


Figure 6. Sample path of the process X with parameters $\mu = 3$, $\sigma = 1$ and the corresponding a posteriori probability process Π with $\lambda = 1/10$ (blue), $\lambda = 1/360$ (magenta), $\lambda = 1/1000$ (cyan). The disorder time θ is indicated by a star, while the optimal stopping times τ_A are indicated by a diamond for $1 - A = 0.03$.

Table 2. Detection times of the simulated trajectory from figure 6 and the theoretical expected detection delay for different values of λ .

λ	τ_A	D
1/10	300.25	1.22
1/360	300.45	2.00
1/1000	300.65	2.22

Note: The disorder time θ equals 300 in this case.

the time series both in the time and in the frequency domain (we omit these results) we can conclude that the time deviation X can be approximated by a Wiener process. Looking at figure 7(a), it seems to be apparent that the observed process X has no zero drift before the disorder time. So we estimate the initial drift $\mu_0 = 2.63 \times 10^{-13}$ and we normalize the observed data by compensating this first trend through the transformation $\tilde{X}_t = X_t - \mu_0 t$. The resulting sample path of the process \tilde{X} is shown in figure 7(c). We now estimate the diffusion coefficient $\sigma = 6.71 \times 10^{-12}$ using Allan variance and the new possible trend $\mu = 1.14 \times 10^{-12}$ from the increments and we apply the detection algorithm with different values of λ .

In figure 8 the a posteriori probability process Π is plotted for different values of λ . The values of τ_A are indicated by a

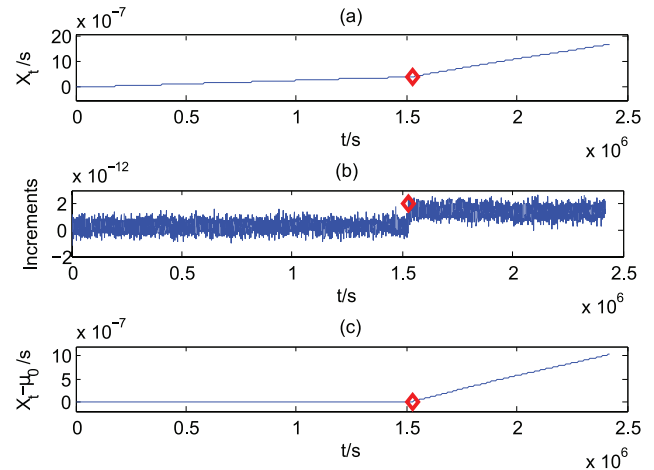


Figure 7. (a) GLONASS data set; (b) Corresponding increments; (c) transformed data $X_t - \mu_0 t$. The optimal stopping time τ_A is indicated by a diamond for $1 - A = 10^{-7}$.

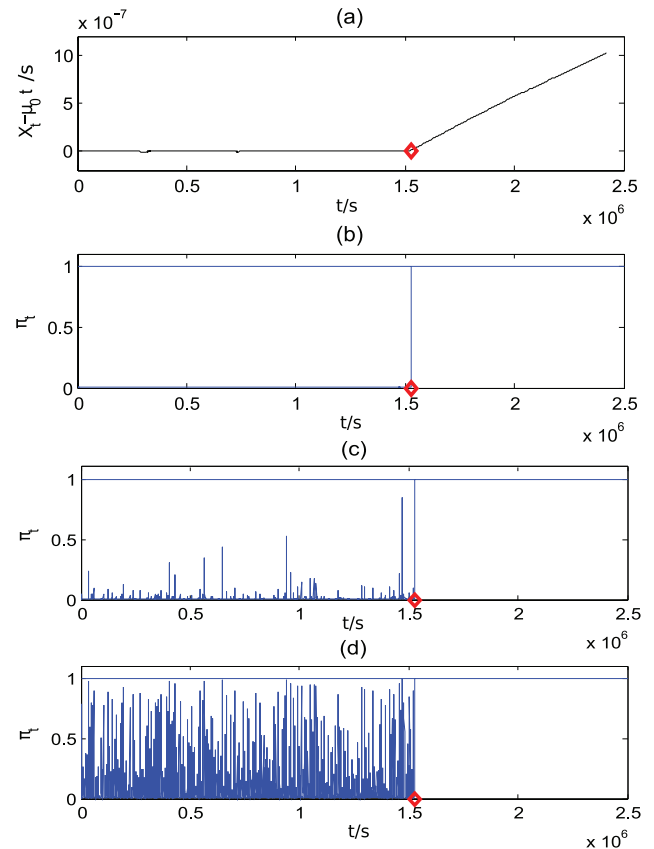
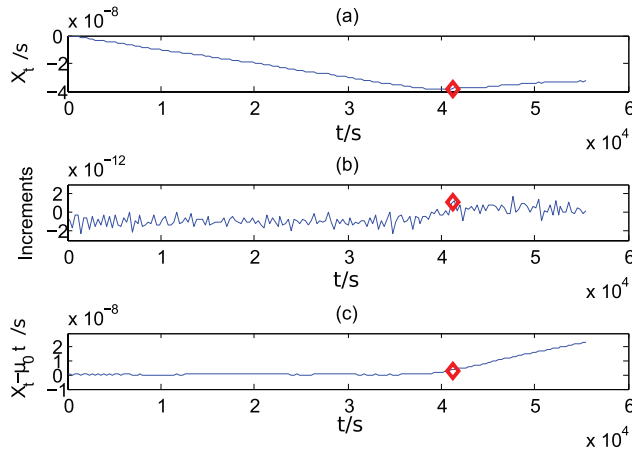


Figure 8. GLONASS data set (a) and the corresponding a posteriori probability process with (b) $\lambda = 1/(3 \times 10^7)$ (one event a year); (c) $\lambda = 1/10^5$ (one event every 28 h); (d) $\lambda = 1/1500$ (one event every 25 min). The optimal stopping times τ_A are indicated by a diamond for $1 - A = 10^{-7}$.

diamond and are very similar. Indeed, it seems that the value of λ does not affect the results significantly. The values of τ_A and of the theoretical expected detection delay for different values of λ are listed in table 3. As λ increases, $\mathbb{E}[\theta] = 1/\lambda$ decreases, which implies that the disorder time is smaller. Indeed, we can see that the a posteriori probability process

Table 3. Detection times of GLONASS data set and the theoretical expected detection delay for different values of λ .

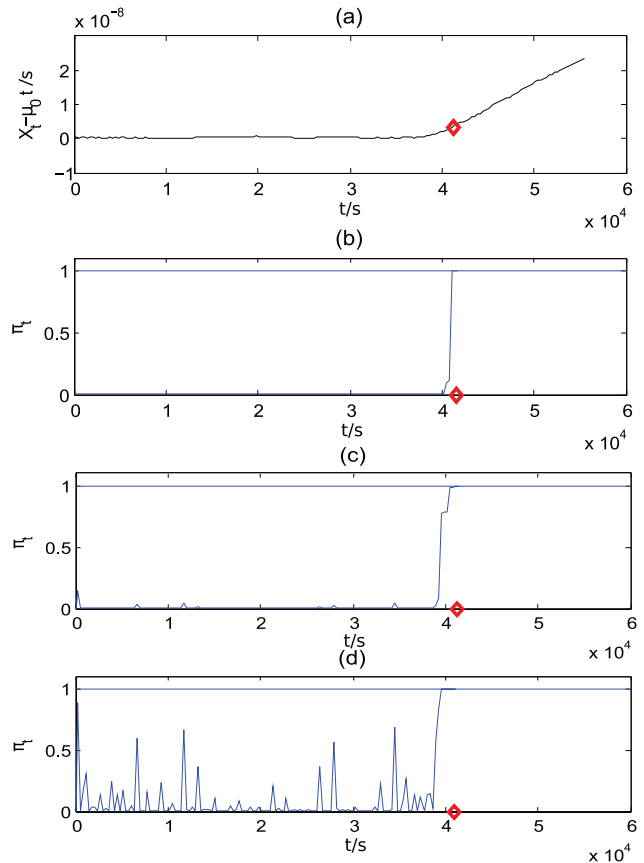
$\mathbb{E}[\theta] = 1/\lambda$ (s)	τ_A (s)	D (s)
$3 \cdot 10^7$	1526 550	1904.94
10^5	1526 250	1509.30
1500	1525 950	1170.71

**Figure 9.** (a) GPS data set; (b) corresponding increments; (c) transformed data $X_t - \mu_0 t$. The optimal stopping time τ_A is indicated by a diamond for $1 - A = 10^{-7}$.

Π is more sensitive to the fluctuations of the process X . The smaller the $1/\lambda$ is in size, the earlier the anomaly is expected (see figure 8(d)). Therefore the algorithm is more eager to declare a change.

The second data set was extracted by a GPS space rubidium clock [9]. The sample size of the data set is 186 and the sampling time is 300 s. The time deviation data and the increments are represented in figures 9(a) and (b). Looking at the increments we note that the change of the drift is not instantaneous, in fact the drift takes some time to stabilise. Even if the analysis of the time series in the time and in the frequency domain does not fully confirm that this data set can be approximated by a Wiener process, we nonetheless apply the detection algorithm. Looking at figure 9(a), it seems to be apparent that the observed process X initially has a negative trend. We estimate the initial drift using the sample mean of the increments $\mu_0 = -1.00 \times 10^{-12}$ and we normalize the observed data by compensating this first trend. The resulting sample path is shown in figure 9(c). We estimate the diffusion coefficient $\sigma = 9.93 \times 10^{-12}$ from the Allan variance and the new possible drift $\mu = 1.38 \times 10^{-12}$ from the increments and we apply the detection algorithm with different values of λ .

In figure 10 the *a posteriori* probability process Π is plotted for different values of λ . The values of τ_A are indicated by a diamond. The values of τ_A and of the theoretical expected detection delay for different values of λ are listed in table 4. Also in this case the optimal stopping times are very similar, and different values of λ do not affect the result. We can conclude that a rough estimate of λ can be sufficient for efficient detection.

**Figure 10.** GPS data set (a) and the corresponding *a posteriori* probability process with (b) $\lambda = 1/(3 \times 10^7)$ (one event a year); (c) $\lambda = 1/55\,800$ (one event every 15 h); (d) $\lambda = 1/1500$ (one event every 25 min). The optimal stopping times τ_A are indicated by a diamond for $1 - A = 10^{-7}$.**Table 4.** Detection times of GPS data set and the theoretical expected detection delay for different values of λ .

$\mathbb{E}[\theta] = 1/\lambda$ (s)	τ_A (s)	D (s)
$3 \cdot 10^7$	41 550	2806.75
55 800	41 250	2152.55
1500	40 950	1677.51

6. Conclusions

The application of the optimal stopping method to the atomic clock data is quite exciting. The method tested on simulated data shows its capability to detect the insurgence of a new drift with a minimal delay even when setting a very low probability of false alarm. A new linear drift in the clock time deviation corresponds to a jump in the frequency values and we know that this can be a serious drawback particularly in GNSS space clocks.

The exposed version of the theorem asks for a considerable knowledge of the process:

- (i) it shall be a Wiener process;
- (ii) the parameters μ , λ , σ , and π are assumed to be known.

Regarding the clock time deviation, the assumption of a Wiener process means that its derivative, i.e. the frequency deviation, is white and this is often the case. Also other cases in which the frequency can be modeled by a Wiener process can be found and the method can be applied as well [7] by interpreting the process X as the frequency deviation. The parameter σ can be easily estimated using the Allan deviation, the parameter $1/\lambda$ indicating the mean occurrence of the anomaly can be guessed as well by the knowledge of the clock, and we have seen that the estimates can also be quite inaccurate and the answers of the method are not that different. The only parameter which really represents an issue is the value of the upcoming drift μ . In fact we know that a frequency jump may occur and that this would lead to the insurgence of a different linear slope on the time deviation, but if we want to apply the detector in real time, we cannot know the actual value of the upcoming new drift. We have seen by simulation that the method can cope with a postulated value of the upcoming drift μ that is not exact, however at least the order of its magnitude needs to be determined in some way.

In addition, since real measurement data, particularly from space clocks, are known to be affected by outliers, missing information, and spurious behaviour, all of which corrupt the pure nature of the Wiener process, a suitable preprocessing of data may therefore be necessary.

Nevertheless we have demonstrated a remarkable performance of this method, owing to the existence of an optimal stopping time, which means that under stated conditions, no better estimation is possible. We are currently working on the extension of the method to a random upcoming drift value and studying the analytical properties of the delay in detection

to offer a more general method with desirable features in the analysis of atomic clock frequency trends.

Acknowledgments

Work partially supported by the University of Torino local grant: POLFRIL013.

References

- [1] Peskir G and Shiryaev A N 2006 *Optimal Stopping and Free-Boundary Problems* (Basel: Birkhäuser)
- [2] Shiryaev A N 1978 *Optimal Stopping Rules* (New York: Springer)
- [3] Galleani L, Sacerdote L, Tavella P and Zucca C 2003 The mathematical model for the atomic clock error: an overview *Metrologia* **40** S257–64
- [4] Zucca C and Tavella P 2015 A mathematical model for the atomic clock error in case of jumps *Metrologia* **52** 514–21
- [5] Atkinson K E 1989 *An Introduction to Numerical Analysis* 2nd edn (New York: Wiley)
- [6] Abramowitz M and Stegun I 1965 *Handbook of Mathematical Functions* (New York: Dover)
- [7] Weaver G L, Jensen J R, Zucca C, Tavella P, Formichella V and Peskir G 2016 Estimation of the dynamics of frequency drift in mature ultra-stable oscillators: a study based on the in-flight performance from New Horizons *Proc. 47th Annual Precise Time and Time Interval Systems and Applications Meeting (Monterey, CA, January 2016)* pp 198–205
- [8] Zucca C and Tavella P 2005 The clock model and its relationship with the Allan and related variances *IEEE Trans. Ultrason. Ferroelectr. Freq. Contr.* **52** 289–96
- [9] Cernigliaro A, Fantino G, Sesia I, Galleani L and Tavella P 2014 Nonstationarities in space clocks: investigations on experimental data *Proc. 28th European Frequency and Time Forum (Neuchâtel, Switzerland, June 2014)* pp 126–9